

Kate Alexandra Kujawa

A Framework for Intelligent
Information Management to Support
Production unit Designers

UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Institut für Technische Produktentwicklung

A Framework for Intelligent Information Management to Support Production unit Designers

Kate Alexandra Kujawa, M. Eng.

Vollständiger Abdruck der von der Fakultät für Luft- und Raumfahrttechnik
der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

Vorsitzender: Univ.-Prof. Dr. rer. nat. Eric Jäggle
Gutachter der Dissertation: 1. Univ.-Prof. Dr.-Ing. Kristin Paetzold
2. Univ.-Prof. Dr.-Ing. Michael Vielhaber
Universität des Saarlandes

Produktentwicklung

Kate Alexandra Kujawa

**A Framework for Intelligent Information Management
to Support Production unit Designers**

Shaker Verlag
Düren 2021

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Zugl.: München, Univ. der Bundeswehr, Diss., 2020

Copyright Shaker Verlag 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 978-3-8440-7819-0

ISSN 1866-1742

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Phone: 0049/2421/99011-0 • Telefax: 0049/2421/99011-9

Internet: www.shaker.de • e-mail: info@shaker.de

Acknowledgements

This dissertation was written during my time as a doctoral student at the *Institute for Technical Product Development at the Bundeswehr University Munich* and while working at *Mercedes-Benz AG*. I found the simultaneous immersion in the world of science and industry made possible by this arrangement to be particularly valuable and broadened my horizon. On the side of science, I would first like to thank my supervisor, Professor Kristin Paetzold, for her guidance and dedicated support. She continuously provided encouragement and insightful feedback. For this, I am extremely grateful. I would also like to acknowledge Professor Michael Vielhaber for taking the position of second-corrector.

I would like to thank my managers Matthias Müller and Patricia Till for the trust they placed in me and the freedom they provided for me to pursue this topic. I thank my colleagues for their support, countless stimulating discussions, and friendship. I would like to express my special thanks to Florian Burlein, Simon Klumpp, Dr. Markus Stäbler, Lissy Langer, Marco Gund, and Uwe Habisreiter. I would particularly like to single out my mentor and good friend Dr. Jakob Weber who encouraged me to start, and was instrumental in helping define the path of my research.

A supporting pillar for a successful doctoral thesis is the collection of students and research participants involved. Many thanks to the students and participants that made this dissertation possible. I would especially like to emphasize Justin Lange who contributed significantly to my work and with whom I hope to continue working.

A special thank you goes to my mother Pamela Kujawa, who spent countless hours proofreading and challenging my research. I would like to express my deepest appreciation to my family and partner, who were by my side each step of the way, demonstrating patience and providing encouragement. They inspired me and made all my academic achievements possible.

Stuttgart, December 10th, 2020.

Abstract

Since the invention of the automobile, the automotive industry has been constantly changing and adapting to overcome obstacles to remain competitive. In recent years, the accelerating pace of innovation and trend toward outsourcing and cross-innovation are intensifying the job of an innovator in this industry. The evolution of the internet and ever increasing advances in techno-science, have resulted in masses of information accessible to production unit designers, who already have highly information-intense jobs, and are now being overwhelmed with this array of information. This particular challenge is known as Information Overload.

These employees spend approximately 20 % of their time obtaining information, this is in part due to the hundreds of internal and external sources accessible by the user. As a result of Information Overload, employees are avoiding academic literature, and tend to prioritize easy to use sources, such as blogs and forums, reducing the quality of information being considered in innovation. Although employees spend significant amounts of time searching for information, nearly 20 % of design work is duplicated because searches are not successful in identifying previous work.

Today, significant advances are being made in Information Retrieval, with search engines constantly advancing and adapting. However, large search engines focus on generic searches for the masses, often with a focus on profit. Current search solutions do not cater to employees with specific search needs and are not tailored for user groups working in a specific domain.

As part of a trilogy of dissertations, with the goal of improving the working methods of production unit designers, this work proposes a novel framework for designing a holistic information assistant to provide optimal information throughout an entire project.

The framework begins by using observation and survey techniques to obtain a current state analysis. Data about existing information practices is gathered, and a Data Flow Diagram is created to illustrate all the information flows and data sources related to a project. To create a personalised tool, user-orientation is key, and in this work, the software development methods Personas and User-Stories provide user context to transform information needs into software requirements. Once information needs and software requirements have been created, solutions meeting the various and specific user needs are conceptualised. The sum of solutions, provides an information platform with all necessary information is called AIIS – an Artificially Intelligent Information Scout. AIIS and the original requirements are validated using Wizard of Oz testing techniques. By building a User Interface and running mock tests, user interaction confirms whether the tool meets the needs of the user. In addition to validation, verification is conducted using experiments and testing to assure functionality.

The proposed framework was applied and validated in a real industry setting, and the approach is outlined in full in this work. Technical solutions for the identified requirements were conceptualised, and AIIS – an Artificially Intelligent Information Scout was created. The validation of AIIS serves to validate the framework. AIIS solutions were found to be 250 % faster in some cases, and applied solutions had a significantly higher average precision for search results, than solutions currently in use. It can be concluded that this framework leads to the development of a holistic information management system designed around user needs.

In this thesis, a framework and technical solutions have been developed and applied, that enable an optimal, holistic, and practical realization of a highly intelligent search assistant, AIIS. The potential of the novel framework and solutions has been demonstrated through various validation and verification techniques. The developed framework was implemented in an automotive assembly unit development branch, however, transferability to other similar high information intense jobs is realistic, and should be investigated in further work

Kurzfassung

Seit der Erfindung des Automobils hat sich die Automobilindustrie ständig verändert und angepasst, um den wachsenden Anforderungen gerecht zu werden und wettbewerbsfähig zu bleiben. In den letzten Jahren haben das zunehmende Innovationstempo und der Trend zu Outsourcing und Cross-Innovation die Arbeit eines Innovators in dieser Branche intensiviert. Die Entwicklung des Internets und immer größere Fortschritte in der Technowissenschaft haben dazu geführt, dass den Entwicklern von Produktionssystemen, die ohnehin schon sehr informationsintensive Arbeitsplätze innehaben, massenhaft Informationen zur Verfügung stehen und nun mit diesem Informationsangebot überhäuft werden. Diese besondere Herausforderung wird als Informationsüberflutung bezeichnet.

Diese Mitarbeiter verbringen etwa 20 % ihrer Zeit mit der Beschaffung von Informationen, was zum Teil auf die Hunderte internen und externen Quellen zurückzuführen ist, auf die der Benutzer Zugriff hat. Infolge der Überlastung mit Informationen vermeiden die Mitarbeiter akademische Literatur und neigen dazu, leicht zugänglichen Quellen wie Blogs und Foren den Vorrang zu geben. Dies verringert die Qualität der Informationen, die bei der Innovation berücksichtigt werden. Obwohl die Mitarbeiter viel Zeit mit der Suche nach Informationen verbringen, wird fast 20 % der Entwicklungsarbeit dupliziert, weil die Suche nach bisherigen Ergebnissen nicht erfolgreich ist.

Heute werden bedeutende Fortschritte im Information Retrieval gemacht, wodurch die Suchmaschinen ständig weiterentwickelt und angepasst werden. Große Suchmaschinen konzentrieren sich jedoch auf die generische Suche für die breite Masse, oft mit dem Fokus auf Profit. Die derzeitigen Suchlösungen richten sich nicht an Mitarbeiter mit spezifischen Suchbedürfnissen und sind nicht auf Benutzergruppen zugeschnitten, die in einem bestimmten Bereich arbeiten.

Im Rahmen einer Trilogie von Dissertationen mit dem Ziel, die Arbeitsmethoden der Entwickler von Produktionseinheiten zu verbessern, schlägt diese Arbeit ein neuartiges Framework für die Gestaltung eines ganzheitlichen Informationsassistenten vor, der während des gesamten Projekts optimale Informationen liefert.

Das Framework beginnt mit dem Einsatz von Beobachtungs- und Erhebungstechniken, um eine Analyse des Ist-Zustandes zu erhalten. Es werden Daten über bestehende Informationspraktiken gesammelt, und es wird ein Datenflussdiagramm erstellt, um alle Informationsflüsse und Datenquellen im Zusammenhang mit einem Projekt zu veranschaulichen. Um ein personalisiertes Werkzeug zu erstellen, ist die Benutzerorientierung der Schlüssel. In dieser Arbeit wandeln die Softwareentwicklungsmethoden Personas und User-Stories, die den Benutzerkontext liefern, die Informationsbedürfnisse in Softwareanforderungen für einen Informationsassistenten

um. Sobald die Informationsbedürfnisse und Softwareanforderungen erstellt sind, werden Lösungen konzipiert, die den verschiedenen und spezifischen Benutzerbedürfnissen entsprechen. Die Summe der Lösungen, die alle notwendigen Informationen für den Benutzer bereitstellen, wird als Informationsassistent bezeichnet. Zur Validierung der Anforderungen werden Wizard of Oz-Testtechniken vorgeschlagen. Durch den Aufbau einer Benutzerschnittstelle und die Durchführung von Mock-Tests bestätigt die Benutzerinteraktion, ob das Werkzeug die Bedürfnisse des Benutzers erfüllt. Zusätzlich zur Validierung wird die Verifizierung anhand von Experimenten und Tests durchgeführt, um die Funktionalität sicherzustellen.

Das vorgeschlagene Framework wurde in einer realen Industrieumgebung angewandt und validiert, und der Ansatz wird in dieser Arbeit vollständig dargelegt. Technische Lösungen für die identifizierten Anforderungen wurden konzeptualisiert, und AIIS - ein künstlich intelligenter Informationsscout - wurde geschaffen. Die Validierung von AIIS dient der Validierung dem in dieser Arbeit vorgeschlagenen Framework. Es wurde festgestellt, dass die AIIS-Lösungen in einigen Fällen um 250 % schneller waren und dass die angewandten Ansätze eine deutlich höhere durchschnittliche Genauigkeit der Suchergebnisse aufwiesen als die derzeit verwendeten Suchmaschinen. Es kann der Schluss gezogen werden, dass dieses Framework die Entwicklung eines ganzheitlichen Informationsmanagementsystems ermöglicht, das auf die Bedürfnisse der Benutzer ausgerichtet ist.

Table of Contents

1	Introduction	8
1.1	Motivation and Problem Description	10
1.2	Goals and Boundary Conditions	11
1.3	Thesis Structure	12
2	State of the Art.....	13
2.1	Data, Information, and Knowledge.....	13
2.2	Design and Development of Production Units	14
2.2.1	Automotive Production and Changeable Assembly Units	15
2.2.2	Product Design and Development	15
2.2.3	Information in Product Design	19
2.3	Information Resources	20
2.3.1	Open Innovation	21
2.3.2	Overview of Sources	24
2.4	Information Overload.....	27
2.4.1	Fundamentals.....	27
2.4.2	Consequences	28
2.4.3	Causes.....	30
2.5	The Science of Search.....	31
2.5.1	Information Retrieval	31
2.5.2	Artificial Intelligence.....	33
2.6	From Knowledge to Data.....	35
2.6.1	Knowledge.....	35
2.6.2	Technology Intelligence – Desired Knowledge	36
2.6.3	Web Search Behaviour Based on Expertise	37
2.6.4	Search Behaviour Based on Information.....	40
2.6.5	Knowledge Discovery Solutions	41

3	Call for Action	44
3.1	Fundamentals and Identifying the Research Gap	44
3.2	Summary of Findings and Research Questions	46
3.3	Structure and Methods	46
4	Information Overload and Information Needs	49
4.1	Effects of Information Overload	49
4.1.1	Data Gathering Methods.....	50
4.1.2	Results	52
4.1.3	Discussion of Results	55
4.2	Information Needs	56
4.2.1	Data Flow Diagram Method.....	56
4.2.2	Results	60
4.2.3	Discussion of Results	62
4.3	Requirements for a Holistic Search Assistant.....	70
4.3.1	Personas	72
4.3.2	User Stories	76
5	Artificially Intelligent Information Scout	80
5.1	Generic Exploration and Monitoring.....	80
5.1.1	Generic Exploration.....	80
5.1.2	Generic Monitoring	85
5.2	Domain-Specific Exploring and Monitoring	86
5.2.1	Information Retrieval with Machine Learning.....	86
5.2.2	Domain-Specific Monitoring.....	90
5.3	Networking Knowledge	91
5.3.1	External Networking.....	92
5.3.2	Internal Networking.....	93
5.4	Extracting Entities.....	95
5.5	Aggregating Sources	97

5.6	Language.....	99
5.7	<i>Adapt!</i> – Adapting Information with AIIS	101
5.7.1	Goals and Evaluation.....	101
5.7.2	Design and Development.....	102
5.7.3	Technical Solution Solving	103
5.8	AIIS.....	103
6	Verification and Validation	105
6.1	Validation with Wizard of Oz Prototyping	105
6.2	Verification through Experiment	110
6.2.1	Integration with <i>Adapt!</i>	111
6.2.2	Saving Time.....	114
6.2.3	Internal Networking.....	116
6.2.4	Integrating Sources.....	116
6.2.5	Learning to Rank	117
6.3	Discussion of Results.....	119
7	Discussion.....	121
7.1	Summary	121
7.2	Limitations	125
7.3	Recommendations for Future Research.....	126
8	Conclusions	128
9	References	130
10	Appendix	143
10.1	Data Flow Diagram.....	143
10.2	User Story	148
10.3	Companies Examined in Market Analysis.....	150
10.4	Input for Humanoid Robot Project	152

Abbreviations

AD	Axiomatic Design
AI	Artificial Intelligence
AIIS	Artificially Intelligent Information Scout
ANP	Analytical Network Process
API	Application Programming Interfaces
DFD	Data Flow Diagram
DIN	Deutsches Institut für Normung
DP	Design Parameter
DSE	Domain-Specific Exploration
DSM	Domain-Specific Monitoring
ETS	External Technology Searching
FR	Functional Requirement
IE	Information Extraction
IR	Information Retrieval
ISO	International Organisation for Standardisation
IT	Information Technology
JSON	JavaScript Object Notation
MAP	Mean Average Precision
ML	Machine Learning
NER	Named Entity Recognition
NLP	Natural Language Processing
OCR	Optical Character Recognition
REST API	Representational State Transfer Application Programming Interfaces
SFViz	Social Friends Visualisation
UML	Unified Modeling Language

Figures

Figure 1.1: Potentially disruptive trends [55].....	8
Figure 1.2: Receptor model of a Production Unit System [171].....	9
Figure 2.1: The Stairs of Knowledge [115].....	13
Figure 2.2: <i>Adapt!</i> – A holistic method for the design and development of changeable assembly systems	16
Figure 2.3: Historic Growth of Cross-Industry Innovation based on [83, 161].....	21
Figure 2.4: Increase in Technology Analysis Publications by [68]	22
Figure 2.5: Scientific Classification of the Terms [132].....	22
Figure 2.6: The four modes of Technology Intelligence [87].....	23
Figure 2.7: Information Retrieval.....	32
Figure 2.8: Information Retrieval System based on [28].....	32
Figure 2.9: Rumsfeld Knowledge Matrix [136].....	36
Figure 2.10: Modes of Technology Intelligence [87]	37
Figure 2.11: Modes of search behaviour by knowledge, based on [76].....	38
Figure 2.12: Process model of information seeking from [76]	39
Figure 2.13: Search Behaviour based on Information [99, 169, 176].....	40
Figure 2.14: Modes of Knowledge Discovery [103].....	42
Figure 3.1: Framework for creating a holistic user-oriented search solution.....	47
Figure 4.1: Cases of duplicated work	53
Figure 4.2: Excerpt from DFD	59
Figure 4.3: Information Matrix	61
Figure 4.4: Data Matrix	62
Figure 4.5: The modes of Information Needs	63
Figure 4.6: Exploration and Domain-Specific Exploration	64
Figure 4.7: Monitoring and Domain-Specific Monitoring.....	67
Figure 4.8: Information Needs vs. Time	69
Figure 4.9: User-oriented requirements	71
Figure 4.10: Persona variables	73
Figure 4.11: Sam’s Persona.....	74
Figure 4.12: Sally’s Persona.....	75
Figure 4.13: Excerpt of User Story	77
Figure 4.14: Current IR set-up	78
Figure 5.1: Methods of Technology Foresight [91].....	81
Figure 5.2: Learning to Rank Framework [69].....	87
Figure 5.3: Machine Learning Pipeline based on [183].....	88

Figure 5.4: Representation of an Edge Bundle Diagram	93
Figure 5.5: Representation of a Graph Database.....	95
Figure 5.6: Information Extraction.....	96
Figure 5.7: External IR.....	98
Figure 5.8: Internal IR.....	98
Figure 5.9 <i>Adapt!</i> integrated with AIIS for optimal Information Support.....	101
Figure 5.10: AIIS Discovery Solutions Matrix	104
Figure 6.1: WOz prototyping set up based on [119].....	106
Figure 6.2: WOz User Interface Input Template.....	107
Figure 6.3: Internal and External Search Results for Domain-Specific Exploring.....	109
Figure 6.4: WOz External Networking Visualisation	110
Figure 6.5: Design Parameters with and without AIIS (Excerpt of AD Hierarchy Tree).....	113
Figure 6.6: Hours required to obtain search results	115
Figure 6.7: Internal Project Networking Graph	116
Figure 6.8: Mean Average Precision comparison	119
Figure 7.1: The Modes of Information Needs.....	122

Tables

Table 1: Information Sources Production Unit Designers Need based on [113, 138].	24
Table 2: Frequency of Sources Used [45].....	55
Table 3: Methods of Technology Foresight [91]	83
Table 4: Auxiliary Methods, based on [91].....	84
Table 5: The Confusion Matrix [20]	88
Table 6: Machine Learning Pipeline Comparison.....	90